

SEQUENCE LISTING

1. GENERAL INFORMATION:

(i) APPLICANT: ROMEO, Tony and WANG, Xin

(ii) TITLE OF INVENTION: METHODS FOR POLYSACCHARIDE ADHESIN  
SYNTHESIS MODULATION

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:

Ridout & Maybee LLP  
19th Floor - 150 Metcalfe Street  
Ottawa, Ontario  
K2P 1P1

(v) COMPUTER-READABLE FORM:

a) COMPUTER: IBM Compatible  
b) OPERATING SYSTEM: MS DOS  
c) SOFTWARE: EditPad

(vi) CURRENT APPLICATION DATA:

a) APPLICATION NUMBER:  
b) FILING DATE:  
c) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

a) APPLICATION NUMBER: US 60/414,352  
b) FILING DATE: 9/30/2002  
c) CLASSIFICATION:

2. INFORMATION FOR SEQ ID NO: 1

(i) SEQUENCE CHARACTERISTICS:

a) LENGTH: 2700  
b) TYPE:  
c) STRANDEDNESS:  
d) TOPOLOGY:

(ii) MOLECULE TYPE: Combined DNA and Amino Acid Sequences

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(vii) IMMEDIATE SOURCE:

(viii) POSITION IN GENOME:

a) CHROMOSOME/SEGMENT:

b) MAP POSITION:  
c) UNITS:

(ix) FEATURE:

a) NAME/KEY:  
b) LOCATION:  
c) IDENTIFICATION METHOD:  
d) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

a) AUTHOR(S):  
b) TITLE:  
c) JOURNAL:  
d) VOLUME:  
e) ISSUE:  
f) PAGE(S):  
g) DATE:  
h) DOCUMENT NUMBER:  
i) FILING DATE:  
j) PUBLICATION DATE:  
k) RELEVANT RESIDUES IN SEQUENCE ID NO:

(xi) SEQUENCE DESCRIPTION: SEQUENCE ID NO: 1

ATGTATTCAAGTAGCAGAAAAAGGTGCCCCGAAAACCAATGGGCTTTGAAACTTCTTACT 300  
M Y S S S R K R C P K T K W A L K L L T  
GCCGCATTTTTAGCAGCGAGTCCCGCGCGAAGAGTGCTGTTAATAACGCCTATGATGCA 360  
A A F L A A S P A A K S A V N N A Y D A  
TTGATTATTGAAGCTCGCAAGGGTAATACTCAGCCAGCTTTGTGTCATGGTTGCACTAAAA  
L I I E A R K G N T Q P A L S W F A L K 420  
TCAGCACTCAGCAATAACCAATTGCTGACTGGTTACAGATTGCCTTATGGGCCGGGCAA  
S A L S N N Q I A D W L Q I A L W A G Q 480  
GATAAACAGGTTATTACCGTTTACAACCGCTACCGTCATCAGCAATTACCAGCGCGTGGT  
D K Q V I T V Y N R Y R H Q Q L P A R G 540  
TATGCAGCTGTCCCGTTCGCTTATCGTAACCTGCAACAATGGCAAACTCGCTTACACTG  
Y A A V A V A Y R N L Q Q W Q N S L T L 600  
TGGCAAAAGGCGCTCTCTCTGGAGCCGCAAAATAAGGATTATCAACGGGGACAAATTTA  
W Q K A L S L E P Q N K D Y Q R G Q I L 660  
ACCTGGCAGATGTGGTCACTATGATACTGCCCTGGTTAACTTAAGCAGCTTAACCT  
T L A D A G H Y D T A L V K L K Q L N S 720  
GGAGCACCGGACAAAGCCAATTTACTCGCAGAACCTATATCTATAAACTGGCGGGGCGT  
G A P D K A N L L A E A Y I Y K L A G R 780  
CATCAGGATGAATTACGGCGGATGACAGAGTCATTACCTGAAAATGCATCTACGCAACAA  
H Q D E L R A M T E S L P E N A S T Q Q 840  
TATCCACAGAATACGTGCAGGCATTACGTAATAATCAACTTGCTGCCGCGATTGACGAT  
Y P T E Y V Q A L R N N Q L A A A I D D 900  
GCCAATTTAACGCCAGATATTTCGCGCTGATATTCATGCCGAACCTGGTCAGACTGTCGTTT  
A N L T P D I R A D I H A E L V R L S F 960  
ATGCCTACGCGCAGTGAAAGTGAACGTTATGCCATTGCCGATCGCGCCCTCGCCCAATAC  
M P T R S E S E R Y A I A D R A L A Q Y 1020  
GCTGCTAGAAATTCGTGGCAGGATAACCCAGACCGCACTGCCCAGTACCAGCGTATT  
A A L E I L W H D N P D R T A Q Y Q R I 1080  
CAGGTTGATCATCTTGGCGGTTATTAACCTCGCGATCGTTATAAGACGTTATTTCTCAC  
Q V D H L G A L L T R D R Y K D V I S H 1140  
TATCAGCGATTAAAAAGACGGGGCAAATTATTCGCCCTGGGGGCAATATTGGGTTGCA  
Y Q R L K K T G Q I I P W G Q Y W V A 1200  
TCGGCTTATCTCAAGATCATCAGCCGAAAAAGCACAGTCAATAATGACCGAGCTCTTT  
S A Y L K D H Q P K K A Q S I M T E L F 1260

TATCACAAGGAGACCATTGCCCGGATTTATCCGATGAAGAACTTGCGGATCTCTTTTAC 1320  
Y H K E T I A P D L S D E E L A D L F Y  
AGCCACCTGGAGAGTGAATAATATCCGGGCGCGCTAACTGTCACCCAACATACCATTAAT 1380  
S H L E S E N Y P G A L T V T Q H T I N  
ACTTCGCCGCTTTCTTCGTTAATGGGCACGCTACGAGCATCCCGAATGATACCTGG 1440  
T S P P F L R L M G T P T S I P N D T W  
TTACAGGGGCATTTCGTTTCTCTCAACCGTAGCAAAATATAGTAATGATCTTCTCAGGCT 1500  
L Q G H S F L S T V A K Y S N D L P Q A  
GAAATGACAGCCAGAGAGCTTGCTTATAACGCACCAGGAAATCAGGGACTGCGCATTGAT 1560  
E M T A R E L A Y N A P G N Q G L R I D  
TACGCGAGTGTGTTACAAGCCGCGGTTGGCCTCGTGCAGCAGAAATGAATTAAAAAA 1620  
Y A S V L Q A R G W P R A A E N E L K K  
GCAGAAGTGATCGAGCCACGTAATATTAATCTGGAGGTTGAACAAGCCTGGACAGCATT 1680  
A E V I E P R N I N L E V E Q A W T A L  
ACGTTACAAGAATGGCAGCAGGAGCTGTCTTAACGCACGATGTTGTGCGAACGTGAACCG 1740  
T L Q E W Q Q A A V L T H D V V E R E P  
CAAGATCCCGCGGTTGTACGATTAACCGTGCAGGTTGATGTACATAATCTTGCAGAGCTT 1800  
Q D P G V V R L K R A V D V H N L A E L  
CGTATCGCTGGCTCAACAGGAATGATGCCGAAGGCCCGGATAGTGGTAAACATGATGTC 1860  
R I A G S T G I D A E G P D S G K H D V  
GACTTAACCACTCGTTTATTCACCAACGCTGAAGGATAACTGGCGCGGTTTGTCTGGA 1920  
D L T T I V Y S P P L K D N W R G F A G  
TTCGGTTATGCCGATGGACAATTTAGCGAAGGAAAGGGATTGTTTCGCGACTGGCTTGGC 1980  
F G Y A D G Q F S E G K G I V R D W L A  
GGTGTGAGTGGCGGTACGTAATATCTGGCTCGAGGCAGAGTACGCTGAACGCGTTTTC 2040  
G V E W R S R N I W L E A E Y A E R V F  
AATCATGAGCATAAACCCGCGCGCTGCTGGCTGGTATGATTTTAAATGATAACTGG 2100  
N H E H K P G A R L S G W Y D F N D N W  
CGTATTGGTTTCGCAACTGGAACGCCTCTCTACCGCGTTCATTACGGGCAATGAAAAAT 2160  
R I G S Q L E R L S H R V P L R A M K N  
GGTGTACAGGCAACAGTGCTCAGGCTTATGTTTCGCTGGTATCAAAATGAGCGGCGTAAG 2220  
G V T G N S A Q A Y V R W Y Q N E R R K  
TACGGTGTCTCCTGGGCTTTCACTGATTTTCCGACAGTAACAGCGTCATGAAGTCTCA 2280  
Y G V S W A F T D F S D S N Q R H E V S  
CTTGAGGGTTCAGGAACGCATCTGGTCTTACCATATTTGATTGTCGATTTCTACCCAGT 2340  
L E G Q E R I W S S P Y L I V D F L P S  
CTGTATTACGAACAAAATACAGAACAGTACCCCATACCAACCTATAAAAAACGTTTC 2400  
L Y Y E Q N T E H D T P Y Y N P I K T F  
GATATTGTTCCGGCATTGAGGCAAGCCATTTGTTATGGCGAAGCTATGAAAATAGCTGG 2460  
D I V P A F E A S H L L W R S Y E N S W  
GAGCAAAATATTCAGCGCAGGTGTTGGTGCCTCCTGGCAAAAACATTATGGCACGGATGTC 2520  
E Q I F S A G V G A S W Q K H Y G T D V  
GTCACCCAACCTCGGCTACGGGCAACGCATTAGTTGGAATGACGTGATTGATGCTGGCGCA 2580  
V T Q L G Y G Q R I S W N D V I D A G A  
ACGCTACGCTGGGAAAAACGACCTTATGACGGTGACAGAGAACACAACCTATACGTTGAA 2640  
T L R W E K R P Y D G D R E H N L Y V E  
TTCGATATGACATTGAGATTTAAGGATAAATATGTTACGTAATGGAATAAATATCTCC 2700  
F D M T F R F \*

2. INFORMATION FOR SEQ ID NO: 2

(i) SEQUENCE CHARACTERISTICS:

- a) LENGTH: 2031
- b) TYPE:
- c) STRANDEDNESS:
- d) TOPOLOGY:

(ii) MOLECULE TYPE: Combined DNA and Amino Acid Sequences

- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- (vii) IMMEDIATE SOURCE:
- (viii) POSITION IN GENOME:
- a) CHROMOSOME/SEGMENT:
- b) MAP POSITION:
- c) UNITS:
- (ix) FEATURE:
- a) NAME/KEY:
- b) LOCATION:
- c) IDENTIFICATION METHOD:
- d) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- a) AUTHOR(S):
- b) TITLE:
- c) JOURNAL:
- d) VOLUME:
- e) ISSUE:
- f) PAGE(S):
- g) DATE:
- h) DOCUMENT NUMBER:
- i) FILING DATE:
- j) PUBLICATION DATE:
- k) RELEVANT RESIDUES IN SEQUENCE ID NO:

(xi) SEQUENCE DESCRIPTION: SEQUENCE ID NO: 2

```

TTAAGGATAAATATGTTACGTAATGGAAATAAATATCTCCTGATGCTGGTGAGTATAATT      60
  M L R N G N K Y L L M L V S I I
ATGCTCACCGCGTGCAATTAGCCAGTCAAGAACATCATTTATACCGCCACAGGATCGCGAA      120
  M L T A C I S Q S R T S F I P P Q D R E
TCTTTACTCGCCGAGCAACCGTGGCCGCATAATGGTTTTGTAGCGATTTTCATGGCATAAC      180
  S L L A E Q P W P H N G F V A I S W H N
GTTGAAGACGAAGCTGCCGACCAGCGTTTTATGTCAGTGCCGACATCAGCACTGCGTGAA      240
  V E D E A A D Q R F M S V R T S A L R E
CAATTTGCCTGGCTGCGCGAGAACGGTTATCAACCGGTCAGTATTGCTCAAATTCGTGAA      300
  Q F A W L R E N G Y Q P V S I A Q I R E
GCACATCGAGGAGGAAAACCGCTACCGGAAAAAGCTGTAGTGCTGACTTTTGATGACGGC      360
  A H R G G K P L P E K A V V L T F D D G
TACCAGAGTTTTTATACCGCGCTCTTCCCAATTCTTCAGGCCTTCAGTGGCCTGCTGTA      420
  Y Q S F Y T R V F P I L Q A F Q W P A V
TGGGCCCCCGTCGGCAGTTGGGTCGATACGCCAGCGGATAAACAAGTAAATTTGGCGAT      480
  W A P V G S W V D T P A D K Q V K F G D
GAGTTGGTCGATCGAGAATATTTGCCACGTGGCAACAAGTGCGAGAAGTTGCGCGTTCC      540
  E L V D R E Y F A T W Q Q V R E V A R S
CGGCTCGTTGAGCTCGCTTCTCATACATGGAATTCTCACTACGGTATTGAGGCTAATGCC      600
  R L V E L A S H T W N S H Y G I Q A N A
ACCGGCAGCTTATTGCCTGTATATGTAAATCGTGCATATTTTACTGACCACGCACGGTAT      660

```

T G S L L P V Y V N R A Y F T D H A R Y	
GAAACCGCAGCAGAATACCGGGAAGAATTCGTCTGGATGCTGTAAAAATGACGGAATAC	720
E T A A E Y R E R I R L D A V K M T E Y	
CTGCGTACAAAGGTTGAGGTAAATCCACACGTTTTTGGTTGGCCTTATGGCGAAGCGAAT	780
L R T K V E V N P H V F V W P Y G E A N	
GGCATAGCGATAGAGGAATTAATAAACTCGGTTATGACATGTTCTTCACCCTTGAATCA	840
G I A I E E L K K L G Y D M F F T L E S	
GGTTTGGCAAATGCGTCGCAATTGGATTCCATTCCGCGGGTATTAATCGCCAATAATCCC	900
G L A N A S Q L D S I P R V L I A N N P	
TCATTAAGAGTTTGCCAGCAAATTATTACCGTACAGGAAAAATCACCACAACGGATA	960
S L K E F A Q Q I I T V Q E K S P Q R I	
ATGCATATCGATCTTGATTACGTTTATGACGAAAACCTCCAGCAAATGGATCGCAATATT	1020
M H I D L D Y V Y D E N L Q Q M D R N I	
GATGTGCTAATTCAGCGGGTGAAAGATATGCAAATATCAACCGTGTATTTGCAGGCATTT	1080
D V L I Q R V K D M Q I S T V Y L Q A F	
GCTGATCCCGATGGTGATGGGCTGGTCAAAGAGGTCTGGTTTCCAAATCGTTTGCTACCA	1140
A D P D G D G L V K E V W F P N R L L P	
ATGAAAGCAGATATTTTTAGTCGGGTTGCCTGGCAATTACGTACCCGCTCAGGTGTAAAC	1200
M K A D I F S R V A W Q L R T R S G V N	
ATCTATGCGTGGATGCCGTTATTAAGCTGGGATTAGATCCACATTAACGCGAGTAAAA	1260
I Y A W M P V L S W D L D P T L T R V K	
TACTTACCAACAGGGGAGAAAAAGCACAAATTCATCCTGAACAATATCACCGTCTCTCT	1320
Y L P T G E K K A Q I H P E Q Y H R L S	
CCTTTGATGACAGAGTCAGAGCACAAAGTTGGCATGTTATATGAAGATCTTGCCGGACAT	1380
P F D D R V R A Q V G M L Y E D L A G H	
GCTGCTTTTGATGGCATATTGTTCCACGATGATGCTTTGCTTTCAGATTATGAAGATGCC	1440
A A F D G I L F H D D A L L S D Y E D A	
AGTGCACCGGCTATCACGGCTTATCAGCAAGCAGGCTTAGCGGGAGTCTGAGCGAAATT	1500
S A P A I T A Y Q Q A G F S G S L S E I	
CGACAAAACCCGGAGCAATTTAAACAGTGGGCCGCTTTAAAGTCGTGCGTTAACTGAC	1560
R Q N P E Q F K Q W A R F K S R A L T D	
TTCACTTTGAAGCTTAGTGCGCGCTAAAAGCCATTGCGGGTCCACATATTAATACTGCA	1620
F T L E L S A R V K A I R G P H I K T A	
CGAAATATTTTGCACCTCCGGTAATACAACCTGAAAGTGAAGCCTGGTTTGCACAGAAT	1680
R N I F A L P V I Q P E S E A W F A Q N	
TATGCTGATTTCTTAAAAAGCTATGACTGGACCGCTATTATGGCTATGCCTTATCTGGAA	1740
Y A D F L K S Y D W T A I M A M P Y L E	
GGTGTGCGAGAAAAATCGGCTGACCAATGGTTAATACAATTGACCAATCAAATTAATAAC	1800
G V A E K S A D Q W L I Q L T N Q I K N	
ATCCCTCAGGCTAAAGACAAATCTATTTAGAATTACAGGCACAAAACCTGGCAGAAAAAT	1860
I P Q A K D K S I L E L Q A Q N W Q K N	
GGTCAGCATCAGGCTATTTCTTCGCAACAACTCGCTCACTGGATGAGCCTATTACAACCTG	1920
G Q H Q A I S S Q Q L A H W M S L L Q L	
AATGGAGTGAAAAACTATGGTTATTATCCCGACAATTTTCTGCATAACCAACCTGAAATA	1980
N G V K N Y G Y Y P D N F L H N Q P E I	
GACCTTATTCGTCTGAGTTTTCAACAGCCTGGTATCCGAAAAATGATTAA	2031
D L I R P E F S T A W Y P K N D ***	
(YCDR STOP CODON)	
(YCDQ START CODON)	

2. INFORMATION FOR SEQ ID NO: 3

(i) SEQUENCE CHARACTERISTICS:

- a) LENGTH: 1560
- b) TYPE:
- c) STRANDEDNESS:
- d) TOPOLOGY:

(ii) MOLECULE TYPE: Combined DNA and Amino Acid Sequences

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(vii) IMMEDIATE SOURCE:

(viii) POSITION IN GENOME:

a) CHROMOSOME/SEGMENT:

b) MAP POSITION:

c) UNITS:

(ix) FEATURE:

a) NAME/KEY:

b) LOCATION:

c) IDENTIFICATION METHOD:

d) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

a) AUTHOR(S):

b) TITLE:

c) JOURNAL:

d) VOLUME:

e) ISSUE:

f) PAGE(S):

g) DATE:

h) DOCUMENT NUMBER:

i) FILING DATE:

j) PUBLICATION DATE:

k) RELEVANT RESIDUES IN SEQUENCE ID NO:

(xi) SEQUENCE DESCRIPTION: SEQUENCE ID NO: 3

AAAATGATTAATCGCATCGTATCGTTTTTTATATTATGTCTGGTGTATGCATACCCCTA 240

M I N R I V S F F I L C L V L C I P L

TGCGTAGCGTACTTTCACTCTGGTGAACGTGATGAGGTTTCGTTTCTTCTGGCCGTTT 300

C V A Y F H S G E L M M R F V F F W P F

TTTATGTCCATTATGTGGATTGTGGCGCGTCTATTTCTGGGTCTATCGTGAACGCCAC 360

F M S I M W I V G G V Y F W V Y R E R H

TGGCCGTGGGGAGAAAACGCACCAGCTCCCCAGTTGAAAGATAATCCGTCTATCTCCATT 420

W P W G E N A P A P Q L K D N P S I S I

ATCATTCCCTGTTTTAATGAGGAGAAAAACGTTGAGGAAACCATACACGCCGCTTAGCA 480

I I P C F N E E K N V E E T I H A A L A

CAGCGTTATGAGAACATTGAAGTTATTGCCGTAAATGACGGTTCAACAGATAAAACCCGT 540

Q R Y E N I E V I A V N D G S T D K T R

GCCATCCTGGATCGCATGGCTGCACAAATCCCCATTGCGGGTCATTTCATCTGGCGCAA 600

A I L D R M A A Q I P H L R V I H L A Q

AACCAGGGGAAAGCCATTGCGCTTAAACCGGAGCTGCCGCGCGAAAAGTGAATATCTG 660

N Q G K A I A L K T G A A A A K S E Y L

GTGTGCATTGATGGCGATGCGTTATTAGACCGCGATGCCGCGGCATATATTGTGGAACCG 720

V C I D G D A L L D R D A A Y I V E P

ATGTTGTACAACCCGCGTGTGGGTGCCGTAACCGGTAATCCTCGTATTCGAACACGTTCT 780

M L Y N P R V G A V T G N P R I R T R S

ACCCTGGTGGGTAAAATTTCAGGTTGGCGAGTATTCCTCAATTATTGGTTTGATCAAGCGA	840
T L V G K I Q V G E Y S S I I G L I K R	
ACCCAGCGTATCTATGGAACGTATTTACCGTTTCCGGTGTTATTGCCGCATTTTCGTGCG	900
T Q R I Y G N V F T V S G V I A A F R R	
AGCGCCCTGGCAGAAGTGGGTTACTGGAGTGACGATATGATCACCGAAGATATTGATATT	960
S A L A E V G Y W S D D M I T E D I D I	
AGCTGGAAGCTGCAGTTGAATCAGTGGACGATTTTTTACGAGCCACGGGCACTGTGCTGG	1020
S W K L Q L N Q W T I F Y E P R A L C W	
ATATTAATGCCTGAAACGTTAAAAGGGCTGTGGAACAGCGCCTGCGCTGGGCTCAGGGC	1080
I L M P E T L K G L W K Q R L R W A Q G	
GGTGCAGAAGTATTCCTCAAAAATATGACAAGGTTGTGGCGCAAAGAAACTTTCGAATG	1140
G A E V F L K N M T R L W R K E N F R M	
TGGCCGCTGTTTTTTGAATACTGCCCTGACGACAATATGGGCCTTCACCTGCCTGGTCGGT	1200
W P L F F E Y C L T T I W A F T C L V G	
TTCATTATTTACGCGTCCAACCTGCCGGTGTACCGTTAAATATTGAATTGACACATATC	1260
F I I Y A V Q L A G V P L N I E L T H I	
GCTGCGACACATACTGCCGGAATATTATTGTGTACGTTATGTTTACTGCAATTTATTGTC	1320
A A T H T A G I L L C T L C L L Q F I V	
AGCCTGATGATCGAGAATCGCTATGAGCATAATCTGACTTCATCGCTTTTCTGGATTATT	1380
S L M I E N R Y E H N L T S S L F W I I	
TGGTTCCCGGTTATTTTCTGGATGCTGAGCCTGGCAACGACATTGGTATCATTTACACGA	1440
W F P V I F W M L S L A T T L V S F T R	
GTCATGTTGATGCCTAAAAAGCAACGCCCGTTGGGTAAGTCCCGATCGCGGGATTCTG	1500
V M L M P K K Q R A R W V S P D R G I L	
AGAGGTTAATATGAACAATTTAATTATTACGACCCGACAATCACCAGTACGTTTACTGGT	1560
RG* M N N L (ycdP)	

2. INFORMATION FOR SEQ ID NO: 4

(i) SEQUENCE CHARACTERISTICS:

- a) LENGTH: 30
- b) TYPE:
- c) STRANDEDNESS:
- d) TOPOLOGY:

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL:

(iv) ANTI-SENSE:

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(vii) IMMEDIATE SOURCE:

(viii) POSITION IN GENOME:

- a) CHROMOSOME/SEGMENT:
- b) MAP POSITION:
- c) UNITS:

(ix) FEATURE:

- a) NAME/KEY:
- b) LOCATION:
- c) IDENTIFICATION METHOD:

d) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

- a) AUTHOR(S):
- b) TITLE:
- c) JOURNAL:
- d) VOLUME:
- e) ISSUE:
- f) PAGE(S):
- g) DATE:
- h) DOCUMENT NUMBER:
- i) FILING DATE:
- j) PUBLICATION DATE:
- k) RELEVANT RESIDUES IN SEQUENCE ID NO:

(xi) SEQUENCE DESCRIPTION: SEQUENCE ID NO: 4

TACAGTTAAG TGTGTTATCG GTGCAGAGCC

30

2. INFORMATION FOR SEQ ID NO: 5

(i) SEQUENCE CHARACTERISTICS:

- a) LENGTH: 31
- b) TYPE:
- c) STRANDEDNESS:
- d) TOPOLOGY:

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL:

(iv) ANTI-SENSE:

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(vii) IMMEDIATE SOURCE:

(viii) POSITION IN GENOME:

- a) CHROMOSOME/SEGMENT:
- b) MAP POSITION:
- c) UNITS:

(ix) FEATURE:

- a) NAME/KEY:
- b) LOCATION:
- c) IDENTIFICATION METHOD:
- d) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

- a) AUTHOR(S):
- b) TITLE:
- c) JOURNAL:
- d) VOLUME:



- e) ISSUE:
- f) PAGE(S):
- g) DATE:
- h) DOCUMENT NUMBER:
- i) FILING DATE:
- j) PUBLICATION DATE:
- k) RELEVANT RESIDUES IN SEQUENCE ID NO:

(xi) SEQUENCE DESCRIPTION: SEQUENCE ID NO: 5

CTCAACGCCT GGCTGATTAA ACCAACTATT C

31

2. INFORMATION FOR SEQ ID NO: 6

(i) SEQUENCE CHARACTERISTICS:

- a) LENGTH: 7500
- b) TYPE:
- c) STRANDEDNESS:
- d) TOPOLOGY:

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL:

(iv) ANTI-SENSE:

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(vii) IMMEDIATE SOURCE:

(viii) POSITION IN GENOME:

- a) CHROMOSOME/SEGMENT:
- b) MAP POSITION:
- c) UNITS:

(ix) FEATURE:

- a) NAME/KEY:
- b) LOCATION:
- c) IDENTIFICATION METHOD:
- d) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

- a) AUTHOR(S):
- b) TITLE:
- c) JOURNAL:
- d) VOLUME:
- e) ISSUE:
- f) PAGE(S):
- g) DATE:
- h) DOCUMENT NUMBER:
- i) FILING DATE:
- j) PUBLICATION DATE:
- k) RELEVANT RESIDUES IN SEQUENCE ID NO:

(xi)

SEQUENCE DESCRIPTION: SEQUENCE ID NO: 6

ATGTATTCAAGTAGCAGAAAAAGGTGCCCCGAAAACCAAATGGGCTTTGAACTTCTTACT  
GCCGCATTTTTTAGCAGCGAGTCCCGCGGCGAAGAGTGCTGTTAATAACGCCTATGATGCA  
TTGATTATTGAAGCTCGCAAGGGTAATACTCAGCCAGCTTTGTCATGGTTTGCCTAAAA  
TCAGCACTCAGCAATAACCAAATTGCTGACTGGTTACAGATTGCCTTATGGGCCGGGCAA  
GATAAACAGGTTATTACCGTTTACAACCGCTACCGTCATCAGCAATTACCAGCGCGTGGT 300  
TATGCAGCTGTGCGCGTCGCTTATCGTAACCTGCAACAATGGCAAACTCGCTTACACTG  
TGGCAAAAGGCGCTCTCTCTGGAGCCGCAAAATAAGGATTATCAACGGGGACAAATTTTA  
ACCCTGGCAGATGCTGGTCACTATGATACTGCGCTGGTTAACTTAAGCAGCTTAACTCT  
GGAGCACCGGACAAAGCCAATTTACTCGCAGAAGCCTATATCTATAAACTGGCGGGGCGT  
CATCAGGATGAATTACGGGCGATGACAGAGTCATTACCTGAAAATGCATCTACGCAACAA 600  
TATCCACAGAATACGTGCAGGCATTACGTAATAATCAACTTGCTGCCGCGATTGACGAT  
GCCAATTTAACGCCAGATATTCGCGCTGATATTCATGCCGAACTGGTCAGACTGTCGTTT  
ATGCCTACGCGCAGTGAAAGTGAACGTTATGCCATTGCCGATCGCGCCCTCGCCCAATAC  
GCTGCATTAGAAATTCTGTGGCACGATAACCCAGACCGCACTGCCCAGTACCAGCGTATT  
CAGGTTGATCATCTTGGCGCGTTATTAACCTCGCGATCGTTATAAAGACGTTATTTCTCAC 900  
TATCAGCGATTAAAAAGACGGGGCAAATTATTCGCGCCTGGGGGCAATATTGGGTTGCA  
TCGGCTTATCTCAAAGATCATCAGCCGAAAAAGCACAGTCAATAATGACCGAGCTCTTT  
TATCACAAGGAGACCATTGCCCGGATTTATCCGATGAAGAACTTGCGGATCTCTTTTAC  
AGCCACCTGGAGAGTGAAAATTATCCGGGCGCGCTAACTGTCACCCAACATACCATTAAT  
ACTTCGCCGCCTTTCTTCGGTTAATGGGCACGCCTACGAGCATCCCGAATGATACCTGG 1200  
TTACAGGGGCATTGCTTTCTCTCAACCGTAGCAAAATATAGTAATGATCTTCTCAGGCT  
GAAATGACAGCCAGAGAGCTTGCTTATAACGCACCAGGAAATCAGGGACTGCGCATTGAT  
TACGCGAGTGTGTTACAAGCCCGCGTTGGCCTCGTGCAGCAGAAAATGAATTAAAAAAA  
GCAGAAGTGATCGAGCCACGTAATATTAATCTGGAGGTTGAACAAGCCTGGACAGCATTA  
ACGTTACAAGAATGGCAGCAGGCAGCTGTCTTAACGCACGATGTTGTGCAACGTGAACCG 1500  
CAAGATCCCGGCGTTGTACGATTAAACGTCGCGTTGATGTACATAATCTTGACAGCTT  
CGTATCGCTGGCTCAACAGGAATTGATGCCGAAGGCCCGGATAGTGGTAACATGATGTC  
GACTTAACCACCATCGTTTATTCACCACCGCTGAAGGATAACTGGCGCGGTTTTGCTGGA  
TTCGGTTATGCCGATGGACAATTTAGCGAAGGAAAAGGATTGTTTCGCGACTGGCTTGCG

GGTGTGAGTGGCGGTCACGTAATATCTGGCTCGAGGCAGAGTACGCTGAACGCGTTTTC 1800  
AATCATGAGCATAAACCCGGCGCGCCTGTCTGGCTGGTATGATTTTAATGATAACTGG  
CGTATTGGTTCGCAACTGGAACGCCTCTCTACCGCGTTCCATTACGGGCAATGAAAAAT  
GGTGTACAGGCAACAGTGCTCAGGCTTATGTTGCTGGTATCAAAATGAGCGGCGTAAG  
TACGGTGTCTCCTGGGCTTTCCTGATTTTTCCGACAGTAACCAGCGTCATGAAGTCTCA  
CTTGAGGGTCAGGAACGCATCTGGTCTTCACCATATTTGATTGTCGATTTCTACCCAGT 2100  
CTGTATTACGAACAAAATACAGAACACGATACCCATACTACAACCCTATAAAAACGTTT  
GATATTGTTCCGGCATTGAGGCAAGCCATTGTTATGGCGAAGCTATGAAATAGCTGG  
GAGCAAATATTACGCGCAGGTGTTGGTGCCTCCTGGCAAAAACATTATGGCACGGATGTC  
GTCACCCAACCTCGGCTACGGGCAACGCATTAGTTGGAATGACGTGATTGATGCTGGCGCA  
ACGCTACGCTGGGAAAACGACCTTATGACGGTGACAGAGAACACAACCTTATACGTTGAA 2400  
TTCGATATGACATTCAGATTTTAAGGATAAATATGTTACGTAATGGAAATAAATATCTCC  
ycdR(+1)  
TGATGCTGGTGAAGTATAATTATGCTCACCAGCGTGCATTAGCCAGTCAAGAACATCATTTA  
TACCGCCACAGGATCGCGAATCTTTACTCGCCGAGCAACCGTGGCCGCATAATGGTTTTG  
TAGCGATTTTCATGGCATAACGTTGAAGACGAAGCTGCCGACCAGCGTTTTATGTCAGTGC  
GGACATCAGCACTGCGTGAACAATTTGCCTGGCTGCGCGAGAACGTTTATCAACCGGTCA 2700  
GTATTGCTCAAATTCGTGAAGCACATCGAGGAGGAAAACCGCTACCGGAAAAGCTGTAG  
TGCTGACTTTTGATGACGGCTACCAGAGTTTTTATACCGCGTCTTCCCAATTCTTCAGG  
CCTTCCAGTGGCCTGCTGTATGGGCCCCGTCGGCAGTTGGGTCGATACGCCAGCGGATA  
AACAAGTAAATTTGGCGATGAGTTGGTTCGATCGAGAATATTTGCCACGTGGCAACAAG  
TGCGAGAAGTTGCGCGTTCCCGGCTCGTTGAGCTCGCTTCTCATACATGGAATTCTCACT 3000  
ACGGTATTCAGGCTAATGCCACCGGCAGCTTATTGCCTGTATATGTAAATCGTGCATATT  
TACTGACCACGCACGGTATGAAACCGCAGCAGAATACCGGGAAGAATTCTGCTGGATG  
723  
CTGTAAAATGACGGAATACCTGCGTACAAAGGTTGAGGTAAATCCACACGTTTTTGT  
GGCCTTATGGCGAAGCGAATGGCATAGCGATAGAGGAATAAAAAACTCGGTTATGACA  
TGTTCTTACCCTTGAATCAGGTTTGGCAAATGCGTCGCAATTGGATTCCATTCCGCGGG 3300  
TATTAATCGCCAATAATCCCTCATTAAAAGAGTTTGCCAGCAAATTATTACCGTACAGG  
AAAAATCACCACAACGGATAATGCATATCGATCTTGATTACGTTTATGACGAAAACCTCC  
AGCAAATGGATCGCAATATTGATGTGCTAATTCAGCGGGTGAAAGATATGCAAATATCAA  
CCGTGTATTGTCAGGCATTTGCTGATCCCGATGGTGATGGGCTGGTCAAAGAGGTCTGGT

TTCCAAATCGTTTGTACCAATGAAAGCAGATATTTTATAGTCGGGTTGCCTGGCAATTAC 3600  
 GTACCCGCTCAGGTGTAAACATCTATGCGTGGATGCCGGTATTAAGCTGGGATTTAGATC  
 CCACATTAACGCGAGTAAAATACTTACCAACAGGGGAGAAAAAGCACAAATTCATCCTG  
 AACAATATCACCGTCTCTCTCCTTTCGATGACAGAGTCAGAGCACAAAGTTGGCATGTTAT  
 ATGAAGATCTTGCCGGACATGTGCTTTTGATGGCATATTGTTCCACGATGATGCTTTGC  
 TTTCAGATTATGAAGATGCCAGTGACCCGGCTATCACGGCTTATCAGCAAGCAGGCTTTA 3900  
 GCGGGAGTCTGAGCGAAATTCGACAAAACCCGGAGCAATTTAAACAGTGGGCCCGCTTTA  
 AAAGTCGTGCGTTAACTGACTTCACTTTAGAACTTAGTGC GCGGTAAAAGCCATTGCGG  
 GTCCACATATTA AAACTGCACGAAATATTTTGCAC TCCGGTAATACAACCTGAAAGTG  
 AAGCCTGGTTTGCACAGAATTATGCTGATTTCTTAAAAGCTATGACTGGACCGCTATTA  
 TGGCTATGCCTTATCTGGAAGGTGTCGAGAAAAATCGGCTGACCAATGGTTAATACAAT 4200  
 TGACCAATCAAATTA AAAACATCCCTCAGGCTAAAGACAAATCTATTTTAGAATTACAGG  
 CACAAAACCTGGCAGAAAAATGGTCAGCATCAGGCTATTTCTTCGCAACAACCTCGCTCACT  
 GGATGAGCCTATTACAACCTGAATGGAGTGAAAAACTATGGTTATTATCCCGACAATTTTC  
 TGCATAACCAACCTGAAATAGACCTTATTCGTCTCTGAGTTTCAACAGCCTGGTATCCGA  
 ycdQ(+1)  
 AAAATGATTAATCGCATCGTATCGTTTTTTATATTATGTCTGGTGTATGCATACCCCTA 4500  
 TGCGTAGCGTACTTTCACTCTGGTGAACCTGATGATGAGGTTGTTTTCTTCTGGCCGTTT  
 TTTATGTCCATTATGTGGATTGTTGGCGGCGTCTATTTCTGGGTCTATCGTGAACGCCAC  
 TGGCCGTGGGGAGAAAACGCACCAGCTCCCCAGTTGAAAGATAATCCGTCTATCTCCATT  
 ATCATTCCCTGTTTTAATGAGGAGAAAAACGTTGAGGAAACCATACACGCCGCTTTAGCA  
 CAGCGTTATGAGAACATTGAAGTTATTGCCGTAAATGACGGTTCAACAGATAAAACCCGT 4800  
 GCCATCCTGGATCGCATGGCTGCACAAATTCCCCATTTGCGGGTCATTCATCTGGCGCAA  
 AACCAGGGGAAAGCCATTGCGCTTAAAACCGGAGCTGCCGCGGCGAAAAGTGAATATCTG  
 GTGTGCATTGATGGCGATGCGTTATTAGACCGCGATGCGGCGGCATATATTGTGGAACCG  
 ATGTTGTACAACCCGCGTGTGGGTGCCGTAACCGGTAATCCTCGTATTGAAACACGTTCT  
 ACCCTGGTGGGTAAAATTCAGGTTGGCGAGTATTCCTCAATTATTGGTTTGATCAAGCGA 5100  
 ACCCAGCGTATCTATGGAAACGTATTTACCGTTTCCGGTGTTATTGCCGCATTTGTCGCG  
 AGCGCCCTGGCAGAAGTGGGTACTGGAGTGACGATATGATCACCGAAGATATTGATATT  
 AGCTGGAAGCTGCAGTTGAATCAGTGGACGATTTTTTACGAGCCACGGGCACTGTGCTGG  
 ATATTAATGCCTGAAACGTTAAAAGGGCTGTGGAAACAGCGCCTGCGCTGGGCTCAGGGC

GGTGCAGAAGTATTCTCAAAAATATGACAAGGTTGTGGCGCAAAGAAAACCTTCGAATG 5400  
TGGCCGCTGTTTTTTGAATACTGCCTGACGACAATATGGGCCTTCACCTGCCTGGTCGGT  
TTCATTATTTACGCAGTCCAACCTGCCGGTGTACCGTTAAATATTGAATTGACACATATC  
GCTGCGACACATACTGCCGGAATATTATTGTGTACGTTATGTTTACTGCAATTTATTGTC  
AGCCTGATGATCGAGAATCGCTATGAGCATAATCTGACTTCATCGCTTTTCTGGATTATT  
TGGTTCCCGGTTATTTTCTGGATGCTGAGCCTGGCAACGACATTGGTATCATTACACGA 5700  
GTCATGTTGATGCCTAAAAAGCAACGCGCCCGTTGGGTAAGTCCCGATCGCGGGATTCTG  
ycdP (+1)  
AGAGGTTAATATGAACAATTTAATTATTACGACCCGACAATCACCAGTACGTTTACTGGT  
TGATTATGTTGCCACAACCATCTTGTGGACATTATTTGCGTTGTTTCATATTCTTATTCGC  
CATGGATCTGCTGACGGGTTATTACTGGCAAAGCGAGGCCAGAAGCCGACTTCAGTTCTA  
TTTTTTGCTGGCAGTGGCGAATGCCGTCGTGTTAATTGTCTGGGCGCTGTACAATAAGCT 6000  
GCGTTTTCAAAAACAGCAGCATCATGCAGCCTACCAATATACGCCGCAAGAATATGCAGA  
GAGCTTAGCAATACCTGATGAGCTCTATCAGCAACTACAAAAAGCCACAGGATGAGCGT  
ACACTTCACCAGCCAGGGGCAAATAAAAATGGTTGTTTCAGAAAAAGCGCTAGTCCGGGC  
ATAAACACCCAAAAACAAAGCCCGGTTGCGCCGGGCTCTGCACCGATAACACACTTAAGT  
TAGGCATGCAGCGTACGTTGGCAAAGTGCCGAACGTACGCAGTCCTCTTTACCGAACCGG 6300  
ACGATCCCAACCATTTTCATCTTCTTCGAAACGTTCCAGCGCGTCACTTAATCCGGAGCAC  
ACGCCGCGAGGCAAATCGCATTGCGTGATATACCGTTGACGATAACCGTCACGTTCTCC  
CCGAGGCGGGTTAAAAACATTTTCATTTGCGCGGCAGTCACATTCTGCGCCTCGTCAAGA  
ATGACGACTGCATTTTCAAAGGTACGTCCACGCATATAGGCGAACGGCGCAATTTCCACC  
TTCCCTATTTCCGGTTCGAGGCAGTACTGCATAAAGGAAGCCCTAAGCGCCGGACCAGC 6600  
ACGTCGTAGACCGGGCGAAAATAGGGAGCAAACCTTTCTGCGATATCTCCAGGTAAGAAG  
CCAAGATCTTCATCGGCTTGACAGAACTGGACGGGTGACGATAATCCTGTGACATCCTTA  
TGTATCAGGGCCTCTGCCGCTTTTGCTGCGCTGATCCAGGTTTTTCCGCACCCGGCTTCG  
CCCGTGGCGAATATCAGCTGCTTACTCTCAATAGCCTTCAGATAGTGAATTGCGCTTCA  
TTTCGCGCGAGGATGGGCGAAGTATCGCGACTGTGCGGGCCATACCAATGGCTTCTACG 6900  
CCGCCCATCTGCACAAGCGAGGTGACCGATTCTTCTTCACGCTGCTTATGGCTGCGCGAA  
TCCCGTCTCAGCACACGTTTTGCCTCGCGACGAGCTTTGATCACTGCTTTTGTCTTCCC  
ATGGAGAGCACCTTGAGTTGTTTGTATTATCACACGCGCCGTTGGCAGCGCGATTATGC  
GCACGAACATCAGAGGGTTGGCTTCCTTGTAAGCCATAGTTTGCTTTTGGATAAAATGCC

GAAAAACGGCTACGCGCACCGTTTACGGCGTCGGTAACACATGAAAAGAAAGGATGAGGT 7200  
TGAAATGCAAAGTGACGAGATGACTACCGGAGGAGAAACTCCGCGAGTGGTGGCGCGT  
TGATTATCTAAAACATGTCCAGTACAGGACGTTACCATCCGCGATCTCCATAGTACTGA  
CTATCACTGCCGGGAAC TTCCGCTGCTACTTAATAAGTACAACAGATCTCGCATTTATTG  
CAACAATATATTTACTTATATTTAACTATAAAACACCATTT CAGTGACATTAGTTTCTAC  
TGGAAAGATGACAGAGTGATGACAGTGATGAAAAAAGCTGTGTGCTTTCAGCAGGATTTG 7500